



86	474.5	3.6	4725	4	US-09-410-551B-24	Sequence 24, Appl	
87	474	3.6	5392	2	US-08-403-852D-1	Sequence 1, Appl	
88	474	3.6	5392	3	US-08-510-646B-1	Sequence 1, Appl	
89	474	3.6	5392	3	US-09-231-818-1	Sequence 1, Appl	
90	474	3.6	5392	4	US-09-635-359B-1	Sequence 1, Appl	
91	472.5	3.6	8438	1	US-07-945-283-1	Sequence 5, Appl	
C 92	472	3.6	36778	3	US-09-105-537-5	Sequence 23, Appl	
C 93	470.5	3.6	4881	4	US-09-863-859-23	Sequence 14931, A	
94	468.5	3.5	6327	4	US-09-252-991A-14931	Sequence 13, Appl	
95	468.5	3.5	11604	4	US-09-385-040-13	Sequence 19, Appl	
C 96	468.5	3.5	38506	3	US-09-320-878-19	Sequence 1, Appl	
C 97	468.5	3.5	38506	4	US-09-141-908-1	Sequence 19, Appl	
C 98	468.5	3.5	38506	4	US-09-657-440-19	Sequence 13, Appl	
99	466.5	3.5	11604	3	US-09-385-028-13	Sequence 13, Appl	
100	466.5	3.5	11604	4	US-09-726-614-13	Sequence 13, Appl	
ALIGNMENTS							
RESULT 1							
US-09-976-594-306							
; Sequence 306, Application US/09976594							
; Patent No. 6673549							
; GENERAL INFORMATION:							
; APPLICANT: Furness, Michael							
; APPLICANT: Buchbinder, Jenny							
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS							
; FILE REFERENCE: PA-0041 US							
; CURRENT APPLICATION NUMBER: US/09/976,594							
; CURRENT FILING DATE: 2001-10-12							
; PRIOR APPLICATION NUMBER: 60/240,409							
; PRIOR FILING DATE: 2000-10-12							
; NUMBER OF SEQ ID NOS: 1143							
; SOFTWARE: PERL Program							
; SEQ ID NO 306							
; LENGTH: 9053							
; TYPE: DNA							
; ORGANISM: Homo sapiens							
; FEATURE:							
; NAME/KEY: misc feature							
; OTHER INFORMATION: Incyte ID No. 6673549 898877.6							
; NAME/KEY: unsure							
; LOCATION: 2006, 2012							
; OTHER INFORMATION: a, t, c, g, or other							
US-09-976-594-306							
Alignment Scores:							
Pred. No.:		0		Length:		9053	
Score:		12922.00		Matches:		2487	
Percent Similarity:		98.26%		Conservative:		2	
Best Local Similarity:		99.18%		Mismatches:		10	
Query Match:		97.78%		Indels:		35	
DB:		4		Gaps:		5	
US-09-522-753-5 (1-2517) x US-09-976-594-306 (1-9053)							
QY	1	MetSerGlySerThrGlnLeuValAlaGlnThrTyrArgAlaThrGluProArgTyrPro	20				
Db	520	ATGTGCGGATCCACACAGCCTGTGCACAGACGTGGAGGGCCACTGAGCCCGCTACCCG	579				
QY	21	ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu	40				
Db	580	CCCCACAGCTTTCTTACCCAGTGCAGATCGCCGCGACGACGACGAGCGTCCGCTC	639				
QY	41	GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln	60				
Db	640	GAGTACACGACCACTCCCGGACATATGCTCCACCTGTGCCCGGCTCCATCATCCAG	699				
QY	61	ProGlnArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln	80				
Db	700	CCCCAGCGCGAGGCGCCCTCTGCTGTGATGTTCCAGCCCGGAGATGACCGTCCAG	759				
QY	81	GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlySerGluMet	100				

Db 1840 CGGAGAGAGTTTCATGACAGCATCCCAAGAACTTTGGCCCTGATCGCATCATCTCCTGGAGAGG 1899  
Qy LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480  
Db 1900 AAGACAGTGGCTGAGTGGTCTCTATTACTTACCTGACTAAGAAGAAATGAGACTATAAG 1959  
Qy SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGln----GlnGlnGlnGln 499  
Db 1960 AGCCTGTGTGACGAGCATATCGCGCGCGGCAAGAGCGCAGCAGCAGCAGCAGCAGCAG 2019  
Qy Gln 519  
Db 2020 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2079  
Qy LysAspGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 539  
Db 2080 AAGATCAGAAGGAG 2139  
Qy AsnAspLysGluAspLeuLeuLysGluLysGluLysGluLysGluLysGluLysGluLys 559  
Db 2140 AACGACAAAGGAAGACCTCTCAAGGAGAAGACAGACACACCTCAGGGGAGGACAACGAC 2199  
Qy GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579  
Db 2200 GAGAAGGAGGCTGTGGCTCCAAAGGCGCAAACTGCCAACAGCCGAGGAGACGCAAA 2259  
Qy GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599  
Db 2260 GGCGCATACCCGCTCAATGGCTAATGAGCCAAAGCAGGAGGAGGAGGAGGAGGAGG 2319  
Qy GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerArgTftrThrGluGluGlu 619  
Db 2320 CAGAGCCCGAGCTGGCTCCATGAGCTGAATGAGAGTTCTGCTGCACAGAAGAAGAA 2379  
Qy MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639  
Db 2380 ATGGAACACGCAAGAAAGTCTCTGGAACACGCGCGCAACTGGTTCGGCCATCGCCCG 2439  
Qy MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659  
Db 2440 ATGGTGGGCTCCAGACTGTGTCAGTGTAAAGACTTCTACTTCAACTACAGAGAGG 2499  
Qy GlnAsnLeuAspGluIleLeuGlnGlnHisLysLysLysMetGluLysGluArgAsnAla 679  
Db 2500 CAGAACTCGATGATGATCTTGCGACGACACAAGCTGAAGATGGAAGAGGAGAGAAACGC 2559  
Qy ArgArgLysLysLysLysAlaProAlaAlaAla--SerGluGluAlaAlaPheProProVa 699  
Db 2560 CGGAGGAAGAGAAAGACG 2619  
Qy lValGluAspGluGluMetGluAlaSerGlyValSerGlyValAsnGluGluMetValGl 719  
Db 2620 GTTGGAGGATGAGAGATGAGAGGCTCGGGCTGAGCGGAAATGAGAGGAGATGTGTGA 2679  
Qy uGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyPr 739  
Db 2680 GGAGGCTGAA----- 2689  
Qy oAlaThrValAsnAsnSerAspThrGluSerIleProSerProHisThrGluAlaAl 759  
Db 2690 -GCCACTGTCAACAACAGCTCAGACACCGAGAGCATCCCTCTCTCACAATGAGGCGC 2748  
Qy aLysAspThrGlyGlnAsnGlyProLysProAlaThrLeuGlyValaAsp--GlyProp 779  
Db 2749 CAGGACACAGGGCAGAAATGGGCCCAAGCCCCCAGCCCTTGGGCGCCGCGAGGGGCCAC 2808  
Qy roProGlyProProThrProPro--ArgArgThrSerArgAlaProIleGluProThrPro 798  
Db 2809 CCCAGGCGCCACCCACCCACCGAGGACATCCCGGCC--CCCACTGAGCCACCCCGC 2867  
Qy AlaSerGluAlaThrGlyAlaProThrProProAlaProProSerProSerAlaPro 818  
Db 2868 GCCTCTGAGCCACCGAGGCCCTTACGCCCCCACCAGACCCCGCATCGCCTCTGCACT 2927

Qy 819 ProProValValProLysGluGluLysGluGluGluThrAlaAlaProProValGlu 838  
Db 2928 CCTCTGTGTCTCCCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2987  
Qy GluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAla 858  
Db 2988 GAGGGGAG 3047  
Qy GluGluProValLysSerGluCysThrGluGluAlaGluGluGluProAlaLysGlyLys 878  
Db 3048 GAGAGCCCTCAAGAGCGAGTGCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3107  
Qy AspAlaGluAlaAlaGluAlaThrAlaGluGluAlaLeuLysAlaGluLysGluGly 898  
Db 3108 GACGCGAGGCGCTGAGGCGCACGCGCGAGGGCGCTCAAGGCAGAGAGAGAGAGAGG 3167  
Qy GlySerGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSer 918  
Db 3168 GGGAGCGGAGGCGCACACAGCCCAAGAGCTCGGGCGCCCCCCCCAGGACGACTCCAGT 3227  
Qy AlaThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeu 938  
Db 3228 GCTACCTGCAGTGCAGACGAGGTGGATGAGCCGAGGGCGGCGACAAAGACCGGCTGCTG 3287  
Qy SerProArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGln 958  
Db 3288 TCCCCAAGGCGCCAGCTCTCTACCCGACTGGCGACCCCCCGGCCAATGCTCTACCCCG 3347  
Qy LysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaIleProProIleGlnVal 978  
Db 3348 AAGCCACTGACCTGAAGAGAGCTGAAGCAGCAGCGCTGCCATCCCCCCCCCATC---GTC 3404  
Qy ThrLysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProPro--Al 998  
Db 3405 ACCAAAGTCCATGAGCCCCCGGAGAGCGAGCTCCCACCAAGAGGAGCTCCCCCAGAC 3464  
Qy aProProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSe 1018  
Db 3465 CCCACCGCCACCAAACTCTGAGCGCGGAGAGCGAGCGCCCTCAGCAGCGCTGCGAGCAG 3524  
Qy rProArgGlySerArgSerProAlaProProAlaAspLysGlu----- 1033  
Db 3525 CCCCCGGGCAAGAGCAGGAGCGCGCGCACCCCGCGCGCAAGAGGAGGAGAGAGAGAG 3584  
Qy -----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTTPTh 1050  
Db 3585 GTTCTTCCAGCCTTCGACGCGAGGCGCCAGAGCTGCTGGGAGACCCCTTCTGTCGAC 3644  
Qy rSerGlyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaPr 1070  
Db 3645 TTCGGGCTGCGCTTCCCGTGGCCCCCGCGTGAAGTGATCAAGGCTCCCGCGCATGCCCC 3704  
Qy oAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAs 1090  
Db 3705 GGACCCCTCAGCCTTCTCTACGCTCCACCTGGTCAACCCACTGCGCCCTGGGCTCCATGA 3764  
Qy pThrAlaArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSe 1110  
Db 3765 CACTGCCGCGCGCTTCTGCGCGCCACCCACCATCTCCAACCCCGCTCCCTCTCATCTC 3824  
Qy rSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSe 1130  
Db 3825 CTCTGCCACGACCCCGAGCGCTCTCGAGAGGCAAAATAGGTGCATCTCCCCAAGGAATGTC 3884  
Qy rValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMe 1150  
Db 3885 GGTCCAGCTCCAGCTCCCGTACTCAGAGCATGCGCAAGGCGCCGCGTGGGCGCTGTGCACCAT 3944  
Qy tGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGl 1170  
Db 3945 GGGGCTGCCCCCTGCCCATGAGCCCCCAAAAGCTGGCACCTTTCAGCGGAGTGAAGCAGGA 4004

QY	1170	uGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGl	1190	Db	5085	GCCTGAGCTGGGCAAGCCGGCGAGAGCCCTTAACTATGAGGACACCGGGGACACCTT	5144
Db	4005	GCAGCTGTCCCAACGGGGCGAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCACAGCCCA	4064	QY	1550	eAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGl	1570
QY	1190	nGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLy	1210	Db	5145	TGCGGGCCACCTCCACAGAGGTTCCGCCGTGACCCACCGGGAGGCCACGCCGCCCTGCA	5204
Db	4065	GGAGGGCTCGCTGAGAGGGACAGCTCTGGGTCTAGTTCCGGGGGGAAGCATCACCAA	4124	QY	1570	nGluGlySerLeuSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProAr	1590
QY	1210	sGlyIleProSerThrArgValProSerAspSerAlaIleThrTyArgGlySerIleTh	1230	Db	5205	GGAGGGCAGCCTTTGCTCCAGCAGGATCCAGAGCGAAAGCTGACGTCGAGCGCTCG	5264
Db	4125	AGGCATTCCAGCACACGGGTGCCCTCGACAGCGCCATCACATACCGGGTCTCCATCAC	4184	QY	1590	gGluIleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerPr	1610
QY	1230	rHisGlyThrProAlaaspValLeuTyLysGlyThrIleThrArgIleIleGluAs	1250	Db	5265	TGAGATGCCCAAGTCCCGCACAGCAGCGTCCCGGAGCACCACCCACATCTGCC	5324
Db	4185	CCACGGCACGCCAGCTGACCTCTGTACAAAGGGCACCATCACAGATCATCGCGAGA	4244	QY	1610	oTyGluHisLeuLeuArgGlyValSerGlyValAspLeuTyArgSerHisIleProLe	1630
QY	1250	pSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTy	1270	Db	5325	CTATGAGCACCTTGCTTCCGGGGCGTGAGTGGGTGGACCTGTATCGCAGCCATCCCC	5384
Db	4245	CAGCCCCAGTCGCTTGGACCGCGCGGGAGGAGACAGCTGCCCAAGGGCCACGTCTCTA	4304	QY	1630	uAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaIaIaTyTy	1650
QY	1270	rGluGlyLysGlyHisValLeuSerTyArgGlyGlyMetSerValThrGlnCysSe	1290	Db	5385	GGCCTTCGACCCACCTCCATACCCCGCGCATCCCTCTGGACGCGCGCTGCTACTA	5444
Db	4305	CGAAGGCAAGAGGGCCACCTTGTCTATGAGGGTGGCATCTCTGTGACCCAGTGCTC	4364	QY	1650	rLeuProArgHisLeuAlaProAsnProThrTyProHisLeuTyProProTyLeuIl	1670
QY	1290	rLysGluaspGlyArgSerSerGlyProProHisGluThrAlaAlaProLysArgTh	1310	Db	5445	CCTGCCCGACACCTTGGCCCCCAACCCACCTTACCCGACCTGTACCCACCTTCTCAT	5504
Db	4365	CAAGGAGGAGCGGAGAGAGCTCAGGACCCCCCATGAGACGGCGGCCCCCAAGCGCAC	4424	QY	1670	eArgGlyTyProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyI	1690
QY	1310	rTyAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLe	1330	Db	5505	CCCGGGCTACCCCGACAGCGGGCGCTGGAGAACCGCGCAGACCATCATCATGACTAC	5564
Db	4425	CTATGACATGATGAGGGCGCGTGGGACAGGACCATCTCTCAGCAGCATCGAAGTCT	4484	QY	1690	eThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLe	1710
QY	1330	uMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHis	1350	Db	5565	CACCTCGCAGCAGATGTCACACACCGCGCCACCCGCCATGGCCGAGCTGATATGCT	5624
Db	4485	CATGGCGCTGCCATCCCGCGGAGCGACACAGCCCCCACCATCTCAAGAGAGCAGACCA	4544	QY	1710	uArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyAlaAlaGlyProArgGl	1730
QY	1350	sIleArgGlySerIleThrGlnGlyIleProArgSerTyValGluAlaGlnGluAspTy	1370	Db	5625	GAGGGGCTCTCGCCCCGAGTCTCTGCTGGTGCACTCAACTACGCTGGGGTCCCGGAG	5684
Db	4545	CATCCGGGGTCCATCACAAAGGATCCCTCGTCTCTACGTGGAGGCACAGGAGGACTA	4604	QY	1730	yIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyTh	1750
QY	1370	rLeuArgArgGluAlaLysLeuLysArgGluGlyThrProProProProProSe	1390	Db	5685	CATCATCGACCTGCCAAGTCCACACCTGCTGTGCTGCTGCCGCCGACACCGAGCAC	5744
Db	4605	CCTGCTCGGAGGCCAAGCTCCTAAAGCGGGAGGCGCGCTCCGCCCCACCGCCCTC	4664	QY	1750	rProAlaThrAlaMetAspArgLeuAlaTyIleuProThrAlaProGlnProPheSerSe	1770
QY	1390	rArgAspLeuThrGluAlaTyLysThrGlnAlaLeuGlyProLeuLysLeuLysProAl	1410	Db	5745	CCCAGCACCGGCATGGACCGCTTGGCTTACCTCCCGACCGCGCCCGACCTTTCAGCAG	5804
Db	4665	ACGGACCTGACCGAGGCTTAAAGACGAGCGCCCTGGGCCCTCTGAAGCTGAAGCGGC	4724	QY	1770	rArgHisSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrTh	1790
QY	1410	aHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProAr	1430	Db	5805	CCGCCACAGAGCTCCCGACTCTCCCGCAGGAGGTCCACACACTTGACAAAACCAACCAC	5864
Db	4725	CCATGAGGGCTGTGTGGCCACGGTGAAGAGCGGGCGCTCCATCCATGAGATCCCGCG	4784	QY	1790	rThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluAr	1810
QY	1430	gGluLeuLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluLysE	1450	Db	5865	CACGTCTCTGTCGAGCGGAGCGAGACCGGATCGAGAGCGGAGCCGGATCGGGAGCG	5924
Db	4785	CGAGAGCTGGGCACACGCGCGAGCTGCCCTCGCCCCGGCGCTCAAGAGGGGCTC	4844	QY	1810	gGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTyArgProGl	1830
QY	1450	rIleThrGlnGlyThrProLeuLysTyAspThrGlyAlaSerThrGlySerLysLy	1470	Db	5925	GGAAAAGTCCATCCTCAGCTCCACACGAGCGTGGAGACGACCCCATCTGGAGACTGG	5984
Db	4845	CATACGCGAGGGCACCCCGCTCAAGTACACACCGCGCGTCCACCATGCTGCCAAAAA	4904	QY	1830	yThrGluGlnSerSerGlySerSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly	1850
QY	1470	HisAspValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLe	1490	Db	5985	TACAGCAGCAGCAGCGGAGCAGCGCAGCAGCGCGGGGGTGGGGGAGCAGCAGCGCG	6044
Db	4905	GCACGACGTACGCTCCCTCATCGGCGAGCCCCCGCGGAGCGTTCACCCCGTGCACCGCT	4964	QY	1850	gProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAl	1870
QY	1490	uAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyArgGluSerLeuLysSe	1510	Db	6045	CCCGCTCCCTCCATGCCACAGCATCTCCCGCATCTCCCTCGGAGCCCGAGATGC	6104
Db	4965	GGATGTGATGGCGGACGCGCGGCACTGGAAAGCTGCTGTACGAGGAGAGCTTGAAGAG	5024	QY	1870	aLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaVa	1890
QY	1510	rArgProGlyThrAlaAspSerSerGlyGlySerIleAlaArgGlyAlaProValIleVa	1530	Db	6105	CCTCCAGCAGAGACCCAGTGTGTTTCAACACACAGGCATGAAGGGTATCATCACCCTGT	6164
Db	5025	CCGGCCAGGAGCGGCAGCAGCTCGGGGGCTCCATTGGCGGCGCGCCCGGTCTATTGT	5084	QY	1890	lGluProSerLysProThrValLeu-----ArgSerThrSerThrSerProVa	1907
QY	1530	lProGluLeuGlyLysProArgGlnSerProLeuThrTyArgGluAspHisGlyAlaProPh	1550				

6165 GAGCCAGCAGCCAGCCAGGCTCTGAGTGGGCGCAGGCTCCACCTCCACCTCTCTCACCCGT 6224 Db  
1907 lArgProAlaAlaThrPheProAlaThrHisCysProLeuGlyGlyThrLeuAspG1 1927 QY  
6225 TCGCCGGCTGCCACATTCACCTGCGCCACCTGCGCCAGTGGCGGCGACCTCTGATGG 6284 Db  
1927 yValTyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaAr 1947 QY  
6285 GGTCTACCTTACCTCATGAGGCGGCTTGTGCTGCCCAAGAGGCGCCCGGGGTGGCCCG 6344 Db  
1947 gProGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSe 1967 QY  
6345 GCAGAGCGGCGCGAGACACCGGCCATGCTCTCGCCAAAGCCCGCCAGCGCGCTC 6404 Db  
1967 rGlyLeuGluProAlaSerProSerLysGlySerGluProArgProLeuValProPr 1987 QY  
6405 CGGGCTGGAGCGCGCTCTCTCCCCCAGCAGAGGCTCGAGCGCCCGCGCTCTAGTGCCTCC 6464 Db  
1987 oValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAl 2007 QY  
6465 TGTCTCTGGCCAGCCACCATCGCCCGCACCTCGCCGCGGAGAACTCTGCACCTCACCACGC 6524 Db  
2007 aSerProAspProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrG1 2027 QY  
6525 CAGCCCGGACCGCGCGCGCACCTGCTCTCGGCTCGGACCGCCGCGCGGAAAGACTCA 6584 Db  
2027 nSerLysProPheSerIleGlnLeuGluLeuArgSerLeuGlyTyrHisGlySerSe 2047 QY  
6585 AAGTAACCTTTTCCATCCAGGACTGGAACTCCGTTCTCTGGGTATACCAGCGCAGCAG 6644 Db  
2047 rTyrSerProGluGlyValGluProValSerProValSerProSerLeuThrHisAs 2067 QY  
6645 CTACAGCCCGGAGAGGGGTGAGCCCGTCAAGGCTGAGCTCAGCCAGTCTGACCCACGA 6704 Db  
2067 pLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGlyGluLeuAr 2087 QY  
6705 CAAGGGGCTCCCCAAGCAGCTGGAAGAGCTCGAAGAGCAGCCTCGAGGGGGAGCTGGG 6764 Db  
2087 gProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuAr 2107 QY  
6765 GCCCAGCAGCAGGCCCCGTGAAGCTTGGCGGGAGCGCCGCCACCTCCACACCTGCG 6824 Db  
2107 gProLeuProGluSerGlnProSerSerProLeuLeuGlnThrAlaProGlyVally 2127 QY  
6825 GCCGCTGCTGAGAGCCAGCCCTGCTCCAGCCGCTGCTCCAGCCGCCCGCCAGGGTCAA 6884 Db  
2127 sGlyHisGlnArgValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTy 2147 QY  
6885 AGGTCACCCAGCGGGTGTCACTGGGCCAGCACATCAGTGAGGTATCATCACAGGACTA 6944 Db  
2147 rThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhePr 2167 QY  
6945 CACCCGGCACACCCAGCAGCTCAGCGCACCTGCGCCCGCCCGCCCTCTACTCTCTTCCC 7004 Db  
2167 oGlyAlaSerCysProValLeuAspLeuArgProProSerAspLeuTyrLeuProPr 2187 QY  
7005 TGGGGCCAGCTGCCCCGTCTGGACCTCCCGCCGCCACCCAGTGACCTCTACCTCCCGCC 7064 Db  
2187 oProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyLysArgSerPr 2207 QY  
7065 CCCGGACCATGTGTGCCCGCGCTGCTCCCGCCAGAGGGGGGAGAGGTCTCC 7124 Db  
2207 oGluProAsnLysThrSerValLeuGlyGlyGluAspGlyIleGluProValSerPr 2227 QY  
7125 AGAGCCAAACAGACGTGCGTCTTGGGTGTGTGAGGAGCGGTATGACCTGTGTCTCCC 7184 Db  
2227 oProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrAr 2247 QY  
7185 ACCGAGGGCATCAGCGAGCCAGGGCACTCCCGAGGTGCTGTACCCGCTGTGTACCG 7244 Db  
2247 gAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerG1 2267 QY  
7245 GGATGGGAAACAGACGAGCCAGCAGGATGGGCTCCAAAGTCTCCAGGCAACACCGACCA 7304 Db

2267 nProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetVallysserLysLY 2287 QY  
7305 GCCCGCAGCCCTTCTCAGCAGAGCTGACGAGAGCACTCCGCCATGTCAAGTCCAGAA 7364 Db  
2287 sGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSe 2307 QY  
7365 GCAAGAGATCAACAGAAAGCTGAACACCCCAACCGGAATGAGCCTGAATACATATCAG 7424 Db  
2307 rGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTy 2327 QY  
7425 CCAGCCTGGAGCGGAGATCTTCAATATGCTCCCGCATCACCAGAACAGCGCTTATGACCTA 7484 Db  
2327 rArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleAr 2347 QY  
7485 TAGAGCCAGCGGTGAGGAACATGCCAGCACCAACATGGGCTGGAGGCCATAATTAG 7544 Db  
2347 gLysAlaLeuMetGlyLysTyrAspGlnTyrGluSerProProLeuSerAlaAsnAl 2367 QY  
7545 AAAGGCATCTCATGGTAAATATGACAGTGGGAAGAGTCCCGCGCTCAGCGCAATGC 7604 Db  
2367 aPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAs 2387 QY  
7605 TTTTAACCTCTGAATGCCAGTGCAGCCTGCCCGCTGCTATGCCCATAAACCGCTGCTGA 7664 Db  
2387 pGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerG1 2407 QY  
7665 CGAGCGGAGTGACCAACACACTCCTCCAGTGGCGGGGAGGAGGCAAGGTCTCTGG 7724 Db  
2407 yArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgPr 2427 QY  
7725 CAGACCCAGCAGCGGAAAGCCAAAGTCCCGCGCCCGCGGCTGCGCATCTGGGACCGGCC 7784 Db  
2427 oProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAs 2447 QY  
7785 ACCTCTGTCTCTCAGTGACCTCGAGGAGAGACTGCAACCGCGGAGCGGCTGCACAA 7844 Db  
2447 nArgValTyrGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLe 2467 QY  
7845 CCGCGTGTGGAGAGCAGGCGCTCTGTCGCGAGGTTCACCGCCATTCCTCTACAACCCCT 7904 Db  
2467 uIleMetArgLeuGlnAlaGlyValMetAlaSerProProProGlyLeuProAlaG1 2487 QY  
7905 GATCATCGGCTCAGCGGGTGTATGCTTCCACACCCCGCGGCTCCCGCGGG 7964 Db  
2487 ySerGlyProLeuAlaGlyProHisAlaTyrAspGluGluProLysProLeuLeuCy 2507 QY  
7965 CAGCGGGCCCTCGCTGGCCCCCACCACCGCTGGGAGAGGAGCCCAAGCCACTGCTCTG 8024 Db  
2507 sSerGlnTyrGluThrLeuSerAspSerGlu 2517 QY  
8025 CTGCGAGTAGAGACACTCTCCGACAGCGAG 8055 Db

## RESULT 2

US-09-144-085-3/c  
; Sequence 3, Application US/09144085  
; Patent No. 6280999  
; GENERAL INFORMATION:  
; APPLICANT: Gustafsson, Claes  
; APPLICANT: Betlach, Mary C.  
; APPLICANT: Ashley, Gary  
; APPLICANT: Julien, Bryan  
; APPLICANT: Ziermann, Rainer  
; TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA  
; FILE OF INVENTION: THEREFOR  
; FILE REFERENCE: 30062-20020.20  
; CURRENT APPLICATION NUMBER: US/09/144, 085  
; CURRENT FILING DATE: 1998-08-31  
; EARLIER APPLICATION NUMBER: 09/010, 809  
; EARLIER FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3

		LENGTH: 33529			
		TYPE: DNA			
		ORGANISM: Sorangium cellulosum			
		US-09-144-085-3			
		Alignment Scores:			
		Pred. No.: 7,28e-21			
		Score: 726.50			
		Conservative: 314			
		Best Local Similarity: 32.36%			
		Indels: 22,08%			
		Query Match: 5.50%			
		Gaps: 3			
		DB: 137			
		US-09-522-753-5 (1-2517) x US-09-144-085-3 (1-33529)			
QY	17	ProArgTyrProPro--HisSerLeuSerTyrProValGlnIleAlaAlaArgThrHisThr	35		
Db	22686	CCGAGGGCTCCAGGACACCATCGCTCCCGGGCTCGACCTTGA-----TCAGCC	22633		
QY	36	AspValGlyLeuLeuGluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerPro	55		
Db	22632	GACCTCGG-----CGCACACTAGCTTCGCTCGCTCGCGCA	22597		
QY	56	GlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGly	75		
Db	22596	GGCTCAGACACACCGCCACATGGCAGCGTCAGCTCTCTATCGAGT---GCCCCAGCA	22540		
QY	76	AsnGluArgSerGln-----GluLeuHis	83		
Db	22539	CGCGAGCGGCTCAGCCCCCAGCTCCACTGACGTACAGCGCCACTTCCAGCCAAA	22480		
QY	84	-----LeuArgPro---GluSerHisSerTyrLeuProGluLeuGly	96		
Db	22479	CAGCCGGGTGGTGTACTCGCTCGCTCCAGACGCTCTCTCTCGCTGCCCGCGG	22420		
QY	97	LysSer-----GluMetGluPheIleGlySerLysArgProArgLeuGluLeu	113		
Db	22419	CGCGAACACCCACTCTCTCAACCCACGGTCGAGATACGGCTCAGCGCTCGCACCTC	22360		
QY	114	ProAsp-----ProLeuLeuArgProSerPro	122		
Db	22359	GTGGAACCGCCGACGGAACACGGGTACACTTCTGTAAGCTCTTCCCATCCGAGCG	22300		
QY	123	LeuLeuAlaThrGlyGlnProAlaGlySerGluAlaSerLeuLysAspArgSerLeuThr	142		
Db	22299	CTGGTGCCTCG-----CCCGGTGACACGAC	22273		
QY	143	GlyLysLeuGluProValSerProProSerProHisThrAspProGluLeuGluLeu	162		
Db	22272	CGCAAGCTT---CCGCGCTCGCTCGCGTCCCGCTCAC-----CACCGCGCATC	22225		
QY	163	ValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArgGlu	182		
Db	22224	CGGCGCGCCCGACACACGCGCGAAGACCCTCCACAGCTCCCGACGGCTCGCGCGCAAG	22165		
QY	183	IleThrMetValGluGlnIleSerLysLeuLysLysGlnGlnGlnLeuGluGlu	202		
Db	22164	CACCGACCGCGGACTCGAAGTGGTCCGCTCGGCGCGCGCGCT-----GCG	22117		
QY	203	GluAlaAlaLysProProGluProGlu-----LysProValSerProProProle	219		
Db	22116	CACCAGCTCCGACACACCCACCTCCCGCTGCTCTTCCAGCCACTTCGCGCCACCGCCCGC	22057		
QY	220	GluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsn-----ArgLysLys	237		
Db	22056	CTGCGCATTCACCGGGCTGTCTCGACCCGACAGCAGCAGCGCAATACCGCAGGGCG	21997		
QY	238	AlaGluAlaAlaHisArgIleLeuGlyLeuGly-----	249		
Db	21996	CGCTCGGCTCGACAGGCTCGCGCGCTCGCGCGCTTCTTCGAGGATGATG	21937		
QY	249	-----	249		

Db	21936	CGCGTTCTCGTCCCGCTGATCCCGAACGACGACACACGCGCGACGACGCGCCGCTTGGC	21877		
QY	250	---ProGlnValGluLeuProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGlu	268		
Db	21876	CCGCCACGACGCGCCTCTTCGACGAGGACGAGCCGCT-----TCCTTCCACCC	21826		
QY	269	AsnIleLysIleAsn-----GlnAlaMetArgLysLysLysLeuLeuTyrPhe	284		
Db	21825	AATGTCGGGCTCGGCTGCTCCGGGTGACGGTCTTCGGCAGCACCTCGTCTGCATCGA	21766		
QY	285	LysArgArg---AsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArgTyrAsp	303		
Db	21765	CAGCACCATCTTGTACGCGCGC-----	21742		
QY	304	GlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArgArgArg	323		
Db	21741	-----CAGCGCGCGCGCGC	21727		
QY	324	AlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArgLys---	342		
Db	21726	CTGCGCATGTCGAGGTTTGA-----CTTCGACGATCC	21694		
QY	343	-----GlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly	360		
Db	21693	CAGGTACAGCGACGCTCGGGCTACGCTCGGTCCAAACACCTCCGCGACGCTCCGCGC	21634		
QY	361	LeuSer---MetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu	379		
Db	21633	CTCGATCGGCTCTCAAGCTCGTCCGTCGCGTCCGCTCCACCGGTGATGCTCTC	21574		
QY	380	SerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProMetLeu	399		
Db	21573	GGCGCACACACACCA-----CGACGACACGCGCTCGCGGATCACCGCTGTGTG	21526		
QY	400	TyrAspAlaAspGlnGlnArg-IleLysPheIleAsnMetAsnGlyLeuMetAlaAspPr	419		
Db	21525	GGCAGGCGCGTTCGCGCGCGTCAGACCTTGGCTGGACCGTCTCTGTTCCGCGAGCC	21466		
QY	419	oMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluTh	439		
Db	21465	ACGGATCACCCCGACGACACGCTCGCGCTCGCGTTCGCGGTGAGA-CAGCCGCTTCAA--	21409		
QY	439	rPheArgGluLysPheMetGlnHisProLysAsnPhely---LeuIleAlaSerPheLe	458		
Db	21408	-----CAACAGCATCCCGCAACCTCGCGCCGACGCGCGCGCGCTCGCGCT	21365		
QY	458	uGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAs	478		
Db	21364	GCACCGAAGACTTTCGACCGACCGTTCGCGGACATCCCTT-----GAGAC	21317		
QY	478	nTyrLysSerLeuValArgArg-----SerTyrArgArgArgGlyLysSerGlnGlnGln	496		
Db	21316	GGCTGAATCCACAAACAGCGGGGTGCTCATCACCGCTACCCCGCGCGCGCGCA	21257		
QY	496	n-----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	509		
Db	21256	GGTGCATTGCGCTCGCGCGCGCGCTGCACCGCAGGTGCAGCGACACGACGACGACG	21197		
QY	510	-----GlnProMetProArgSerSerGlnGln	518		
Db	21196	AGCACGCGGTGTCCACCGTCATCGTGGCGCTTCGAGCGCCAG-----CGGTAGGCCACCC	21140		
QY	518	uGluLysAspGluLysGluLysGluAlaGluLysGluGluLysProGluVa	538		
Db	21139	GGCGCGAGAGATGCTCGCGCGCTCCCGTCCCGGTACCGTCCAGTCTCTTCAGAT	21080		
QY	538	lGluAsnAspLysGluAspLeuLeuLys---GluLysThrAsp---AspThrSerGlyGln	556		
Db	21079	CGACAGTATGAGACACCGCTAGTCCGAGCCCATCGAACCCAGGTACACTCCGCTGCGCG	21020		
QY	556	uAspAsnAspGluLysGluAlaValAlaSerLys-----	567		
Db	21019	TCCGCTCAGCGCGCGCGCGCGCGCTCAATGCTTCCACGCGCGCTTCCA	20960		

```
QY 568 -----GlyArglyThrAlaAsnSerG1 575
Db 20959 GCACACCGATGCTCGGATCCATCATGCTCGCGCTCGCGGGGATA-TCCCGAAGAAG 20901
QY 575 nGlyArglySerGlyArglyThrArgSerMetAlaAsnGluAlaAsnSerGluGluAl 595
Db 20900 TTCCGTCGAAACAGGTTCGATGTCCTCCGAAACACCCACCTTCGCGCGGTAGCTTCCT 20841
QY 595 allerThrProGlnGlnSer-----AlaGluLeuAlaSerMetG1 608
Db 20840 CGCGCTCCGATCAGGTCGTAGACCTCGAACCCGTCACCGCGCTGGAGGCGCTTCG 20781
QY 608 uLeuAsnGluSerArgTrpThrGluGluMetGluThrAlaLysLysGlyLeuLe 628
Db 20780 ATCGCGTCTTCCCTCGGCCAAGAGCGCGAGTAGTCGTCGCGCGTCGCAAGCGCTCCC 20721
QY 628 uGluHisGly-ArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrValSerG 648
Db 20720 GGCAGCGCGGACGCCATCGACACGATCGGATCGGCTCGCTTTTCCCTGACGACGACGC 20661
QY 648 lncYsLysAsnPheTyPheAsnTrpLysLysArgGlnAsnLeu----- 662
Db 20660 ACCGCGCGCGCTTCTTCACTGACGCCCGCGAAGACGCTGCTTTCAGCAGCGCTCT 20601
QY 663 -----AspGluIleLeuGlnGlnHisLysLeuLysM 673
Db 20600 GCGATCGCCCGCGCGTGGGTAGTCAAGACACGAGCTCGCGGCAAGGACGTCTCGGTA 20541
QY 673 etGluLysGluArgAsnAlaArgArgLysLysLys----- 685
Db 20540 CGGACGTGAGCGGTTGCGCAGCGCCACCGCATCATCGAGTGCAGCGCCAGCTCCTTG 20481
QY 686 -----AlaProAlaAlaAlaSerGluGluAlaAlaPheProp 698
Db 20480 GCACCTGATCGGTCGCACAGCTCGCGCGCTGAGCGCTGAGCGCGCAACCTCGC 20421
QY 698 roValGluAlaAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluMetV 718
Db 20420 CCGCACACGCTCGACGAGCGCAFTTACGCGCTCGCGCTCGCGCAGCGCGGAGCGCT 20361
QY 718 alGluGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerG 738
Db 20360 CGCGAGCGCGGCGCTCTCGCTGTCGCGAGGACGCTTGCGAAG-----CTCG 20307
QY 738 lProAlaThrValAsnAsnSerSerAspThrGluSerIleProSer-ProHisThrGlu 757
Db 20306 GGCGCAACAGCGACGAAACAGCCCGGCGCTCGCGCTGAGCTCCAGCCCGCTGCA 20247
QY 758 AlaAlaLysAspThrGly----- 763
Db 20246 GCTGCGGAGATCCAGGTGCATCGGACAGGCTCGCTTCAGGCGGTGAGAGCGCGCGT 20187
QY 764 -----GlnAsnGlyProLysProAlaThrLeu----- 773
Db 20186 CCAGCAGAGCGAGGCGCTTTCACCGGCTCGGACGAAACCAATTCGCGCTCATGCGCG 20127
QY 773 ----- 773
Db 20126 ACAGCTCGGCTCGCCACAGGTGCGCTGATCCCAACGCGCGCTGTGTCCAGNAGCCCC 20067
QY 774 -----Gly 774
Db 20066 ACGCCAGGCTCGTCGCGCGAGCGCGCGCGCGCGGTGCGCGCGAGCGGTGCGAGGA 20007
QY 775 AlaAspGlyProPro-----ProGlyProPro-----ThrProProArg 787
Db 20006 ACGTGTTTCGCGCGCGGTAGTGTCTCTGGCGCGCGCGCGCGCAAGAGTACCAGCGCGCGCAGC 19947
QY 788 ArgThrSerArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThr 807
Db 19946 AGAACAGCAGACGCGCGAGATCCAGCTCCCGCGTCACTGCTGTCAGGTGACGCGCC 19887
```

---

```
QY 808 ProProProAlaProProSerPro-----SerAlaProProProValValProLys 824
Db 19886 CGTCCACCTTCGCGCGCGCACCGCGCGAGAGGCGCTCGCGCGGTGAGCAGCAGC 19827
QY 825 GluGluLysGluGluThrAlaAlaAlaProProValGluGluGluGluGln--- 843
Db 19826 CGTGTGTCGACGACCGCGCGCGCGGTGTCAGCGACCGCGCTTCAGCGGCGCGCGTGCAGTC 19767
QY 844 LysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLys 863
Db 19766 CGCGCACACACCGCGCGACCTCTCCGCTTCGACAGCTCGCAGCAGCACCGCTCACCG 19707
QY 864 SerGluCysThrGluGluAlaGluGluProAlaLysGlyLysAspAlaGluAlaAla 883
Db 19706 TCTCGCGCGCGAGCTCTCGAGTGACTGACGAGCTCGCGCGCGCGCGCGCTCCAGCC 19647
QY 884 GluAlaThrAlaGluGlyAlaLeuLysAlaGluLysGluGlySerGlyArgAla 903
Db 19646 CGC-----GACGCGAGCTGACGACCAAGGTGCTCACCCGT--- 19611
QY 904 ThrThrAlaLysSerSerGlyAlaProGlnAspSerSerAlaThrCysSerAla 923
Db 19610 -----GCGCGCGCACCGAGTGGCGCGATCGCTGACCCAGCTCCCTGTCGCCGCG 19557
QY 924 AspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeu-----Leu 938
Db 19556 TCACACGACCGTCCCGCGAGGTCAGCTCGCGGCTCTCGGGTGAGCTTCTTCGCTG 19497
QY 939 SerPro---ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerPro 957
Db 19496 CCCTCTCGCGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19437
QY 958 GlnLysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaIleProIleGln 977
Db 19436 GCTCCGCGCG-----CCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19398
QY 978 ValThrLysValHisGluProProArgGluAspAla----- 989
Db 19397 CGTCCAGCGGCTCGGTCCCGACATCGATCAAGCGCGCGCGCGCGCGCGCGCGCTG 19338
QY 990 AlaProThrLysProAlaProAla-----ProProProProGlnAsnLeuGluPro 1007
Db 19337 GCCTCGCGCGCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19278
QY 1008 GluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAla 1027
Db 19277 GCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19218
QY 1028 ProAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProPro 1047
Db 19217 CCGACAGCCACGCTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 19173
QY 1048 CysTrpThrSerGlyLeuProPheProValProProArgGluValIleLysAlaSerPro 1067
Db 19172 -----CGTGGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19122
QY 1068 HisAlaProAspProSerAlaPheSerTyAlaProProGly----- 1081
Db 19121 CGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19062
QY 1082 -----HisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArg 1098
Db 19061 GATCGGCAATCGCTCGCGCTCAGGGCTT-----CGCGCGCGCGCGCGCGCG 19014
QY 1099 ProProThrLysSerAsnProProProLeuLysSerSerAlaLysHisProSerValLeu 1118
Db 19013 CTCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18966
QY 1119 GluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTySer 1138
Db 18965 CCGCGA-----GCTCCACAGGCTGGAAGCTCACCTGATACATATGTGTGGCCAC 18915
QY 1139 GluHisAlaLysAlaProValGlyProValThrMetGlyLeu----- 1152
```



17979	GCTGAAACCGCTACGTCGGCAGCT	-----CCAGCAGCAGC-----CGC	17941
1435	HisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGly	1454	
17940	CGTGGCCCGCAGCAGCAGCCGCTGTCAGTCCACT	-----CGT	17905
1455	ThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArg	1474	
17904	GCCCTCTGCAGTCAGCTCCCGCAGCTCTGTACAGTGGGCAACACCCAC	-----CTTCGT	17848
1475	SerLeuIleGlySerProGlyArgThrPheProValHisProLeuAspValMetAla	1494	
17847	CGCGCTGCAGGCTCCCGCAGCACA	-----CCCCCTGGCCTCCCGCAGCCGCTCTCTCCA	17734
1495	AspAlaArgAlaLeuGluArgAla	-----CysTyrGluGlySerLeuLysSerArgProGly	1513
17793	GCAGCATTCGCAGCAGCCGGTGGCGCTCACCTCCAGAACACACACCGCTCTCTCCA	17734	
1514	ThrAlaSerSerGlyGlySerIleAla	-----Arg	1529
17733	GCAGCTTCGACAGCGCTCGCTCCAGCGCAGCGTCCGCCAAGTTCGGCAGCAGTACT	17674	
1529	eValProGluLeuGlyLysProArgGln	-----SerProLeuThrTyrGluAspHisGly	1547
17673	CGCGCTGCAGCGCTCGCGCGCAGACCTCTCTGTCTACCGTCGAGTAAACGCGAGCT	17614	
1548	-----AlaProPheAlaGlyHis	1553	
17613	CGCTCGCTTCGGCTCGAGCGCAGCAGCTTCGCTCTAGCTCGGCGCAGCGCATCCA	17554	
1554	-----LeuProArgGlySerProValThrMetArgGluProThrProArgLeuG	1570	
17553	TGTGCGCGCTGTGCGAGCGCTAGTCAGCTTCACCTTCGCGGAGAACACACCTTCGCGG	17494	
1570	lnGluGlySerLeuSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProA	1590	
17493	TCAGCTCCACCTCAGCCCGCTCCACCGCTTCACAGTCCACCGCAGCACACCGT	17435	
1590	rgGluIleAlaLysSer	1595	
17434	GACGTGTTACCGCGCTATGCAAGCGCTCCCGCTCCCGCGCATGGCTCTCTGCACC	17375	
1596	-----ProHisSerThrValProGluHisHis	1608	
17374	TCCGACACCGCGCTCGACGAGCATCATCGCCCATCCCGACCGCTCGCGCAGCGCC	17315	
1608	leSerProTyrGluHisLeuArgGlyVal	1624	
17314	TGGTTCGCGCGCAGCTATCCCGCTCCCTCCGCGAGCAGCAGCGCTCCGCGACACACC	17255	
1624	yrArgSerHisIleProLeuAlaPheAspPro	1635	
17254	GCGCGCGACACTCGCTCGCTGT	17196	
1635	hrSerIleProArgGlyIleProLeuAspAlaAlaIleTyrTyrLeuProArgHisL	1655	
17195	GCGCGACCGCGCGCAGACCCAGCATCGGAGCAGCGCGGCTGCACCGCTCCAC	17136	
1655	euaProAsnProThrTyrPro	1673	
17135	CGCTTCGAGCAGCGCTCTCTCCCGCTATCTCCGCGAGCAGCAGCAGCAGCACA	17076	
1673	yrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerG	1693	
17075	GCGCTTCACGCGCGCAGCGCTCATCGCAGCGCTGCA	17038	
1693	lnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyL	1713	
17037	-----CCGCTCCGGAACCGCGCTCTCTCTCCAGCAGCGCGCCGCC	16995	
1713	euSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleAla	1733	
16994	CATCCCGCGCAGCTGGCTCCCTT	16947	





TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1922 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-372-652-9

Alignment Scores:  
Pred. No.: 9.5e-22 Length: 1922  
Score: 718.00 Matches: 212  
Percent Similarity: 48.81% Conservative: 76  
Best Local Similarity: 35.93% Mismatches: 214  
Query Match: 5.43% Indels: 88  
DB: 2 Gaps: 25

US-09-522-753-5 (1-2517) x US-08-372-652-9 (1-1922)

QY 1975 ProSerLysGlySerGluProArgProLeuVal----- 1985  
DB 209 CCAAGTGGCAAGCCCGAGCTCATGCTCAGTAGTGTATCTGAGGCTGGGAAAGATAAA 268  
QY 1986 -----ProProValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsn----- 2001  
DB 269 GGGCTCTCTCCAAATCCAGATATAGGAGAGCTTAAGACCCGAGGGAAGACTTACCAT 328  
QY 2002 LeuAlaProHisAlaSerProAspProProAlaProProAlaSerAlaSerAspPro 2021  
DB 329 ACTGCAGCTAACTTCTATAGACGTGACCATCACCCGCGCAATTCCTCGACAAAGATCG 388  
QY 2022 HisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluArgSerLeu 2041  
DB 389 AGGGAACGTGGCTCTCAAGATTTCAGACTTCT-----AGTAGCTTG 430  
QY 2042 GlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSerSer 2061  
DB 431 TCTTCTCACAGGTATGAAACGGCTAGTGATGCTTGAAGTGAAGTCCGCCAGCTCA 490  
QY 2062 ProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHis 2081  
DB 491 CCTGCACACCCCGAGGAAAGCCACAGGCTATACGCCAGCATGTTAAGCAAAATCAA 550  
QY 2082 LeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2101  
DB 551 GCAGAAATGAGTCCACTCGACAGTATGAGGTCCA----- 586  
QY 2102 HisLeuProHisLeuArg-----ProLeuProGluSerGlnProSerSer 2116  
DB 587 ---CTGCATCATTTATCGGTCCCGAGGAAATCACCATCTCCACAGCAACAGCCACCTG 643  
QY 2117 SerProLeuLeuGlnThr-----AlaProGlyValLysGlyHisGlnArgValThr 2134  
DB 644 CCCCCATCTTCCAGTCAGAGGAATGGACAGGTGCCAGGACCCCATCGACTGATCACA 703  
QY 2135 LeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisProGlnGln 2154  
DB 704 CTTGCTGACCATCTGTCAAATATACACAAGATTTTGTAGAAAT-----CAAGTT 757  
QY 2155 LeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAla-----SerCysPro 2172  
DB 758 CCCTCGCAGCTTCTACTTCTCATATTCCAAACTTCACCATCTGCTTTGTCATCCACCT 817  
QY 2173 ValLeuAspLeuArgProProSerAspLeuTyrLeuProProAsp----- 2189  
DB 818 GTA-----AGAACTAAACTCAAGCCGCTACAGCCAGCAATCAGACTCTCAGACT 868  
QY 2190 -----HisGlyAlaProAla---ArgGlySerProHis-----SerGlu 2201

14968 GCAGATGATGTGCGAAAGCAAGTGCAGGCAAGTTGCTTCGCCAGCAAGCGGAGC 14909  
2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIlelleArgLysAlaLeuMetGly 2352  
14908 CGATTGGGACTTCGACGCCCATG----- 14885  
2353 LysTyrAspGlnTrpGluLysSerProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372  
14884 -----AGCGAATCCATTTCGAGCTCCCTCAAGGCTGGCGGTAGGAATC 14840  
2373 AlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHis 2392  
14839 GCGTCA-----CCGCGCTAGAGCCGAGCACCGCGGAGCCTCGCGCGGCACAAAG 14789  
2393 ThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGlyArgProSerSerArg 2412  
14788 TCGATGAGCGAGTC-----AGGCGCTCCGCTCCGGCAGCGCGCGAGGCGC 14741  
2413 LysAlaLysSerProAlaProGlyLeuAlaSerGlyAsp-----ArgProProSerVal 2430  
14740 TCGCGGAGGTGACGCTCTCTTCGCGGCGAGGAGCGCTTGGCGAGCCCGGGCGC 14681  
2431 SerSerValHisSerGluGly-----AspCys 2439  
14680 ACCAGCGCGGAAGCAGCGCGGAGCTCGCGCTGGACTCGAGCCCGAGCTGC 14621  
2440 AsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySer 2459  
14620 GCGAGATCCAGGTGCACTGGGACCAAGCTCGCTTCAGACGTGAGAGCGCGCGTCCAGC 14561  
2460 ThrProPheProTyrAsn-ProLeu-----IleMe 2469  
14560 AGAGCGAGGCGCTCGCGCCGATATCGGCAGAGCCCTCGCGCTGATGCGCGACAGT 14501  
2469 tArgLeuGlnAlaGlyValMetAlaSerProPro-----ProGlyLeuProAla-- 2486  
14500 TCGGCTCGCGCCAGGTGCGCTGTCATCCCGACCGCGCTTGGCGGCCCAACCCCGAGCC 14441  
2487 -GlySerGlyProLeuAla 2492  
14440 AGGCTCGTCCCGCGAGCC 14422

SULT 3  
-08-372-652-9  
Sequence 9, Application US/08372652  
Patent No. 5932699  
GENERAL INFORMATION:  
APPLICANT: Moore, David  
APPLICANT: Seol, Wongi  
APPLICANT: Choi, Hwang-sik  
TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING  
TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street, Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/372,652  
FILING DATE: 13-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/POCKET NUMBER: 00786/246001

14968 GCATGATGGTGGTGGAAAGCAAGTGGCAGGCAAGTTCCTCCGACCGAAGCGGAGC 14909

2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleAlaArgLysAlaLeuMetGly 2352

14908 CGATTGGGACTTCGAGCGGCATG----- 14885

2353 LysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372

14884 -----AGCGAATCCATTCGAGCTCCCTCAAGGGCTGGGCCGTAGGAATC 14840

2373 AlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHis 2392

14839 GCGTCA-----CGCGCTGAAGCCGAGCAGCCGCGGCGACCTCGGCCCGCACAAAG 14789

2393 ThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGlyArgProSerSerArg 2412

14788 TCGATGAGCGAGCTC-----AGCGCTCCGCTCCGCGCAGCGCGAGGCGC 14741

2413 LysAlaLysSerProAlaProGlyLeuAlaSerGlyAsp-----ArgProProSerVal 2430

14740 TCGCGAGCGCTCGACGCTCTTCTCGCGCAGAGAGCGCTTGCAGCGCCGCGGCGC 14681

2431 SerSerValHisSerGluGly-----AspCys 2439

14680 ACCAGCGCGAGACGCGCGGCGAGCTCGCGCTGGACTCCAGCCCGCTCGAGCTGC 14621

2440 AsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySer 2459

14620 GCGAGATCCAGGTGCACCTGGGACCGAGCGCTCGCTTCAGAGCGTGAGAGCGCGGCGTCCAGC 14561

2460 ThrProPheProTyrAsn-ProLeu-----IleMe 2469

14560 AGACGAGGCGCTCGTCCAGCGATATCGGACGAGCGCTCGCGCGCTGATGCGCGAGT 14501

2469 tArgLeuGlnAlaGlyValMetAlaSerProProPro-----ProGlyLeuProAla-- 2486

14500 TCGGCTCGCCAGGTGGCTGTATCCCGAGCGCGCTTGGCGGCCACCAACCCCGCAGCC 14441

2487 -GlySerGlyProLeuAla 2492

14440 AGGCTCGTCCGCGGAGCC 14422

RESULT 3

US-08-372-652-9

Sequence 9, Application US/08372652

Patent No. 5932699

GENERAL INFORMATION:

APPLICANT: Moore, David

APPLICANT: Seol, Wongi

APPLICANT: Choi, Hwang-Sik

TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING

TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street, Suite 3100

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/372,652

FILING DATE: 13-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 00786/246001





[illegible]







**This Page Blank (uspto)**